

DOCUMENT RESUME

ED 254 562

TM 850 177

AUTHOR Hansen, Joe B.
TITLE Changes in Evaluation Practice and Policy under Compensatory Education.
PUB DATE Mar 85
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (69th, Chicago, IL, March 31-April 4, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Information Analyses (070)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Compensatory Education; Data Analysis; Educational Assessment; *Education Service Centers; Elementary Secondary Education; *Evaluation Methods; Evaluation Utilization; Federal Programs; Measurement Techniques; *Program Evaluation; Program Improvement; School Districts; State Agencies; Testing
IDENTIFIERS Evaluation Research; *Technical Assistance Centers; *Title I Evaluation and Reporting System

ABSTRACT

Major changes have occurred in compensatory education evaluation practices and policies since the introduction of the Title I Evaluation and Reporting System (TIERS) and the Technical Assistance Centers (TAC) in 1976. This paper has three purposes: (1) to discuss these changes at the state and local levels; (2) to illustrate these changes with examples from TAC experiences and case studies from state and local districts; and (3) to identify unmet evaluation improvement needs and foreshadow future improvement. Excerpts from TAC studies are used to illustrate evaluation improvements in four major areas: (1) testing and measurement practices; (2) data quality control; (3) reporting procedures; (4) computer technology; (5) dissemination of evaluation results and feedback to local educational agencies (LEAS). TAC's will continue to have a strong role in addressing unmet evaluation needs in program improvement, coordination of Chapter 1 with other school programs and services, better utilization of TAC services by larger school districts, and more evaluation of migrant and other special Chapter 1 programs. (BS)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED254562

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. B. Hansen

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

CHANGES IN EVALUATION PRACTICE AND POLICY
UNDER COMPENSATORY EDUCATION

Joe B. Hansen

Northwest Regional Educational Laboratory

Portland, Oregon

Paper presented at the 1985 annual meeting of the American Educational
Research Association, Chicago, Illinois, April 1-5, 1985.

TM 850 177

CHANGES IN EVALUATION PRACTICE AND POLICY
UNDER COMPENSATORY EDUCATION

Joe B. Hansen
Northwest Regional Educational Laboratory
Portland, Oregon

Introduction

Many changes have occurred in compensatory education evaluation since the advent of the Title I Evaluation and Reporting System (TIERS) in the mid 1970s. Some of these changes have been the subject of national studies and have been reported by others at various times. In a USOE funded study of the implementation of the TIERS (Reisner et al., 1982), changes were reported in a wide variety of evaluation and evaluation related practices at state and local level. For example, Alkin found that Title I evaluations were being used by SEAs to assist in the recognition of problems at the local level and to inform programmatic intervention. In the same major study, Boruch et al. found that: 1) SEA and LEA evaluators were more knowledgeable about issues involved in estimating effectiveness than before TIERS; and, 2) that there had been an increase in the number and quality of studies and analyses by LEAs in efforts to improve educational programs. These same investigators found major improvements in needs assessments by SEAs and LEAs away from subjective, unsubstantiated claims of need to more rigorous systematic processes. Improvements were also noted in student selection, program planning and specification of objectives, achievement testing-data quality, and capabilities, knowledgeability and independence of evaluation staff. Stonehill and Anderson (1982) examined six areas of improvement in the evaluation of compensatory education programs under ESEA Title I and ECIA Chapter 1. They used SEA and LEA case studies to illustrate improvements in program improvement evaluation, testing procedures, needs assessment, quality control (QC) systems, sustained effects and summer drop off and finally, exemplary practices.

Collectively, these studies indicate that improvements have occurred across a broad range of evaluation practices and policies at state and local level. Data quality has improved substantially, evaluation is more broadly accepted and the general level of sophistication of the Chapter 1 evaluator is much higher than it was in 1976. These improvements however, are anything but uniform and systematic. Rather, they reflect the diversity of the state departments of education and the school districts in which they have occurred. In their efforts to bring about changes in practice, the Technical Assistance Centers (TACs) have adapted to this diversity and developed approaches and methods that respond to the needs and conditions they have encountered. Support for this diversity is found in the observation — "that states vary considerably in their ability and willingness to evaluate" (Boruch et al., 1981). These authors note two classes of SEAs, those who implement TIERS fully and those more cautious or less able to adopt the conventional reporting model (p.76). Boruch and colleagues are cautious about attributing

changes to the TIERS, the TAC system or the federal guidance provided through these systems. They cite U.S. Department of Education's (ED) stress on evaluation and reporting and to some extent technical assistance as responsible for the changes.

Reisner et. al noted several problems in compensatory education evaluation that remain to some degree today. They are instability of the TIERS due to changes in the law; estimates of project impact that overestimate achievement gains (LEAs should shift to annual testing cycles); evaluation models have limited value for small districts; TIERS alone is insufficient for meeting local data needs; relative high cost of TAC service; and varying evaluation capabilities of SEAs (services should vary according to SEA needs.)

This paper has three purposes: 1) to discuss some of the major changes that have occurred in compensatory education evaluation practice and policy at state and local level since the TIERS and TAC services were introduced; 2) to illustrate these changes through examples drawn from TAC experience and case studies of selected states and local districts; and, 3) to identify some of the unmet needs for improving evaluation use and practice and foreshadow future improvements. The approach taken here is to draw upon the literature, TAC experiences and case studies to identify and discuss changes in such evaluation related activities as testing practices, data analysis and error screening, uses of technology, reporting procedures and the use of evaluation for program improvement. Although the case studies focus on the SEA, changes at LEA level will also be examined. Implications for future TAC services will be discussed.

The case studies used here were developed by various TAC staff from across the country. The SEA case descriptions were based on four questions of interest: 1) What was the condition or situation at the time TAC services were first introduced (e.g. metric used, reporting procedure, quality of data, etc.); 2) What specific changes took place in the evaluation and reporting process and what were the roles, respectively, of the TAC and SEA staff; 3) What is the current practice; 4) Does the SEA have an evaluation unit and if so, does the unit have responsibility for the Chapter 1 evaluation? These questions were not used in a rigid format, but rather to provide some semblance of structure and guidance for development of the case studies. Information for the case studies was drawn from a variety of sources including personal experience of the TAC staff, TAC reports and documents and interaction with TAC clientele. The LEA examples include information on district characteristics and changes in evaluation and reporting practice on such things as test selection and testing practices, data analysis and reporting, use of evaluation and program improvement evaluation.

While the case studies were developed by TAC staff and are therefore subject to bias due to the perspective of the observers, they were based on TAC records, personal experience, discussions or correspondence with SEA staff. Statements attributing change to the TAC influence were primarily those of SEA staff, speaking through their TAC representative.

State Level Changes

Seventeen SEA case studies were collected from TAC staff who agreed to assist. Six were from region four, the western states; four were from region three, southwest; three were from region three, southeast; and four from region two, midwest. Since the guidelines mentioned above were not rigidly adhered to there was considerable variability evident in the way each was written. Nonetheless, taken as a whole they reveal a substantial amount of commonality in the major changes observed. They also define a range of TAC roles and SEA-TAC relationships and provide a sense of the critical issues and most difficult implementation problems.

An easily forgotten but critical area of change is the advent and mandated adoption of the normal curve equivalent (NCE) as the metric for aggregating gains across projects. As the case studies illustrate, the pre NCE metric of choice was the grade equivalent (GE) with its strong intuitive appeal and elastic scale properties. It would have seemed miraculous to many who were trying to influence the transition to NCEs in 1976 to have been afforded a glimpse of the widespread acceptability and adoption of the NCE in 1984. Improved measurement practices, including choice of metric for aggregation, test selection, test administration, and better, more informed use of achievement test information are areas of change most often reported. Closely related changes in improved data quality, improved reporting procedures and applications of computer technology follow closely. Dissemination of evaluation results and feedback to LEAs are improvements not mentioned in previous studies. The following excerpts from the TAC case studies illustrate some of the more pronounced changes. The excerpts are clustered by the emphases that emerge from the individual case studies, beginning with measurement issues, followed by reporting and data quality control, data use and dissemination.

Changes in testing practices were among the first and the most visible changes to occur as illustrated by this case study excerpt from state M:

"... pre TIERS evaluation data consisted of a variety of metrics, including grade equivalents. Evaluations were conducted separately for each LEA with little emphasis on aggregating data at the SEA level. Two major changes have occurred since TIERS was adopted, in 1977. First, the evaluation tests were changed in all districts. New tests were selected to obtain a better match between teaching and testing. Second, testing procedures were modified so that tests were given on norm dates and out-of-level testing was used when necessary."

In state F major improvements in testing led to other improvements:

"When TAC services began, GE scores were widely used and irregular practices were common in the timing and administration of tests. Feedback to LEAs was erratic. Tests were being used which had poor and/or outdated norms. The evaluation is now being conducted much more appropriately. For example, tests are being selected to better match the curriculum and are being administered within two weeks of the norm dates in the vast majority of cases. Tests with more recent

norms and generally higher quality are more frequently being used. ...This year over 95 percent of the districts reported data useable for the state report -- the highest figure ever. Reports are now being regularly sent to the districts. Those reports are in an increasingly useable format."

State D provides an example of a quick transition to NCE scores.

"In 1975-76 Title I student achievement data was reported as gain or loss in grade-equivalent by categorical increments of 'cognitive' gains or losses; the increments being 0.1-0.5, 0.6-1.1, 1.2-2.0, and over 2.0...Reporting student achievement gains in terms of NCEs became a requirement ...for the 1977-78 report -- prior to being federally mandated by the education amendments of 1978."

Major changes in testing practices, from extreme variety to a commonly used test statewide, occurred in state H where:

"SEA personnel report that the condition of evaluation at the time TAC services were first introduced was poor. Districts were reporting data based on a variety of different norm and criterion referenced tests. About three fourths of all districts reported and about 95 percent of those reported scores in grade equivalents. The remaining districts reported in percentiles... There are only two districts this year which are unable to report test scores. The current practice...is for districts to report CTBS scores in NCEs."

In many states the average number of performance objectives attained provided the basic measure of effectiveness for the compensatory education programs. State G provides the following example of this once common practice and the changes that have occurred:

"...Districts were required to establish performance objectives and use an objective measure of their choice to determine the extent to which the objectives were attained. Districts often used test data of one type or another, but a wide variety of indicators were reported. Test scores, when they were used, were reported in units corresponding to the phrasing of objectives, and were often grade equivalent scores...Non parametric tests, including the Chi-square and Kolgorov-Smirnov were used to test whether the discrepancy between achieved and intended attainment was significant."

"The introduction of TAC services resulted in immediate and dramatic changes in evaluation and reporting procedures in... TAC staff assisted in adapting the federal requirements into a data collection form for state use....NCE scores replaced the plethora of measures previously used. a TAC developed program to aggregate, analyze and report evaluation data using SPSS was adopted and replaced the non-parametric measures formerly employed. The program provided quality control checks as well as a large assortment of analyses. Within two years, achievement measures reported in NCE scores were being collected from the great majority of districts... Currently,...model A is used uniformly with approximately 75 percent of districts testing on a fall to spring schedule. ...An increasing

number of districts aggregate their achievement data using microcomputer systems developed by the TAC, and some 140 individuals have received training in microcomputer data base applications....At the state level, the old SPSS system is in the process of being replaced by a new TAC developed SPSS-PC system, allowing the evaluation unit to control directly the process of data entry and report preparation."

State J presents another example where the mastery of objectives approach was used. Other measurement problems, including a proliferation of tests, were also present and reporting was a problem for the SEA:

"Projects were evaluated in terms of the extent to which performance objectives were met, e.g. 75 percent of the participants in the reading program will read at grade level by the end of the year. Another typical project objective might be '75 percent of reading students will show one month's gain for each month spent in the program, as measured by the (specific test).' A scale of 1 to 5 was used to measure the extent the project goals were met. G. E. scores were the common for all districts except (the largest district in the state). Raw scores were averaged by grade level in many LEAs... The choice of tests was left to the LEA. At one time, 43 tests were reported by the SEA level evaluator, to be in use by the LEAs. ...Reporting consisted of summarizing LEA outcomes in broad general terms and the annual report took several months for the SEA evaluator to compile. Feedback reports to the LEAs on how they compared to other LEAs in the state or to the national average, were non-existent."

Quality Control in the data collection, analysis and reporting process is one of the major improvements in compensatory education evaluation. Approaches vary widely across states and in many instances have become increasingly sophisticated as the microcomputer has become a widely available tool. Mainframe computers also continue to play a major role. For example, in state C:

"...The state department credits the TAC with working with LEAs to introduce the concept of quality to the evaluation process. The SEA staff established policy, and where possible, provided technical assistance to districts in addressing issues of quality control, data interpretation and the like, but the SEA utilized the expertise of the TAC to occupy the roles of educator and change agent in improving the equality of evaluation reporting....In the last two years... the language of both the SEA and the TAC has changed to encompass not only the interpretation of evaluation results but also the use of evaluation results to determine program improvement efforts.

Quality control and improved reporting via computer characterizes the major change area in state N, where:

"The quality of the districts' annual Chapter 1 reports is verified by the SEA to assure the accuracy and consistency of the information provided. Until recently all achievement results were calculated manually. Totals, percentages, and averages were recalculated and test score conversions were reconverted, all by hand with the assistance of only an adding machine....Presently the numbers provided in the districts' reports are verified by the SEA's computer, using a system developed with the assistance of the...Chapter 1 Evaluation Technical Assistance Center....The system develops a unique evaluation report for each district and combines the information into statewide summaries for use in reporting to the State Board of Education and to the federal Education Department. The system is also capable of receiving data electronically from a district rather than through the traditional transfer of paper and subsequent manual entry...The system was designed to, and has been successful in reducing the reporting burden on the districts and the SEA."

A similar situation prevailed in state P, where:

"...Because they had no one on staff familiar with data analysis, the SEA did not know how to process their data...TAC assisted by using the SAS package that was already available to the SEA on the ...mainframe to 'quality control' and aggregate their results. They assisted the SEA in developing a data quality control procedure that is utilized by their Chapter 1 monitors....TAC assisted the SEA in developing a computer generated feedback report for each LEA, comparing the building, district, state and national performances by grade and subject matter. TAC assisted...in getting on a schedule whereby evaluation is presented to LEAs at the beginning of the school year."

In state B improvements in the testing process and choice of metric were followed by an awareness of data quality control as a major issue as is shown in this example:

"...grade equivalents were aggregated on the basis of different achievement tests. It was impossible to do any type of comparative analysis of gain scores throughout the state. In many instances the only testing that took place in the schools was done in the Title I programs....School districts were not trained in administering, scoring and interpreting the (test) results. There were no quality control measures implemented at either the district or state level and consequently the ability to accurately interpret the evaluation data was very limited. Outside consultants were relied upon by school districts to assist in the evaluation as the SEA did not have the level of sophistication to provide much assistance. During this time the SEA did not provide feedback to the districts in terms of achievement gains or quality control measures....It was not until the ...TACs that any consideration was given to the issue of quality data. The TAC initially developed checklists to be used by the SEA to review the data of districts submitted to the state. With the use of the checklists and feedback to the districts came the need and interest on the part of both the SEA and the LEA to upgrade the

skills of the district personnel. The TACs then became involved in a large scale training program that was implemented by the SEA. Quality control was the central issue of the training effort....Currently, the SEA collects NCE scores from each school and aggregates the data across grade levels, subject area and testing intervals."

State O provides another example of how improved testing leads to improvements in quality control and reporting:

"Two major changes occurred following the advent of TIERS and TAC. First, test administration has greatly improved in terms of adhering to norm dates and the use of out-of-level testing. Second, through quality control measures, high quality evaluation data are now made available for Chapter 1. Evaluation results are shared with the LEAs through the publication of a widely disseminated annual report...More recently, TAC efforts have resulted in an increased awareness on the part of both the SEA and the LEAs, of the importance of using evaluation data for program improvement. Various program improvement activities have since been initiated in a number of LEAs."

Not all quality control procedures are dependent upon the use of a computer. In many states a simpler manual approach preceded efforts to computerize the process, as is shown by this example from state D.

"Services provided by TAC such as the Quality Control Survey two years ago are helpful. Individual district achievement data was examined. Several districts reporting extreme NCE gains were selected, agreed to participate, brought their student achievement data in to regional sites and were assisted by TAC personnel in spotting sources of error in their evaluation procedures. Not adhering to testing within acceptable range of norming dates and ...chance level scores accounted for many of the inflated gains. As a result, these districts are now aware of proper procedures and their student achievement evaluation is more meaningful."

Changes in reporting processes have followed other improvements as shown by some of the previous examples. Major improvements in the length and usefulness of LEA reports, improved forms, SEA technical assistance and the use of process evaluation characterize the progress made in state A.

"The annual evaluation reports were incredibly long, about 50 pages, and included a lot of unnecessary data such as attendance and membership figures that were not used by the SEA. The length of the report grew each time the federal government requested new information, but was not deleted later. The section of the report on process evaluation was just a narrative and waste of time... No technical assistance was provided to the LEA in evaluation. The SEA people who traveled to monitor programs had no knowledge of how to look at data and therefore ignored the evaluation component. No feedback was given to the districts regarding the correctness of their evaluation."

In this same state the TAC receives much of the credit for improvements in reporting as well as other evaluations reforms:

"The TAC helped the SEA to cut the evaluation reports down to a manageable and useful size ...The TAC taught the SEA and LEA people about process evaluation. The TAC staff suggested formats and workshops to train in this area. The TAC helped the SEA to devise an application for process evaluation that could be used consistently across the LEAs. The change from a straight narrative cut the reading time from one hour to five minutes per LEA report, which provided better information in this shorter and more effective format. ...The SEA people attribute their learning about good evaluation practices to the TAC training, and they in turn have taught other SEA staff outside the Chapter 1 area about evaluation. The SEA people are now able to monitor LEA data; they know what to look for in quality control issues...Because evaluation has been upgraded too much, attention is now going to using the information for program improvement efforts."

Change came slowly in state J, the state in which 43 tests had been reported and the number of objectives mastered was a common measure of effectiveness. It began with piloting of the TIERS in 1977 and culminated in widespread improvements:

"The TIERS was phased into operation ...over a two year period from 1976 to 1978. In 1976-77 pilot districts were sought to implement the system, ...volunteer LEAs received followup training from the TAC in model implementation." The number of pilot districts grew slowly from one in 1976-77 to 25 in '77-'78, 40 in '78-'79 and full implementation in 1979-80. In about 1981 the TAC began assisting the SEA in producing a summary booklet of achievement results for feeding back information to the LEAs. The booklet was introduced at the annual statewide conference, in a TAC workshop. With the assistance of TAC and the application of microcomputer technology the booklet has been successively improved each year until it now includes state and national three year trend data by grade level and major program emphasis, against which each LEA can compare its gains."

Changes in the dissemination and use of evaluation results are a final link in a cycle of improvements that began with changes in testing.

An example of improved dissemination and use of data comes from state D, where progress has occurred in knowledge and understanding regarding data interpretation and use:

"Administration uses data as an indicator of how well the program is functioning. Data is used to help identify both programs (which are) obtaining good results and those needing improvement. Some of the local districts use state summary data as a standard with which to compare their district's results."

Boruch et al. noted improvements in front-end evaluation activities, one of which is student selection. This is the focus of change in state K,

where with TAC assistance major improvements have been made in developing high quality selection plans at the district level, based on multiple data sources. Excerpts from the case study illustrate not only the degree of change accomplished, but the close teamwork of the TAC and the SEA:

"In the summer of 1981, the Technical Assistance Center, working with the ...Department of Education, identified several problems at the district level in student selection procedures. (State K) and its TAC implemented a joint effort in overcoming the deficiencies found. The State revised its application form to request much more specific information from school systems. Applications submitted by the systems were reviewed much more thoroughly than in prior years. The state worked directly with the systems to improve their plans. TAC was asked by the SEA to visit LEAs with exceptional problems; and to assist the districts in formulating acceptable and feasible plans. ...TAC provided training in student selection through workshops available to all school systems...The State of ... has progressively increased its standard for selection plans....Evidence ...can be found by comparing recent selection plans with those of preceding years. Districts have provided much greater detail. They have in most cases, used at least two different sources of information reflecting academic achievement. Data used are substantially more objective, specifically in areas where teachers are asked to provide ratings of student performance."

Some LEA Examples

Turnbull (1982), examined technical assistance and program implementation in compensatory education at the local district level, with an emphasis on the relationship between technical assistance and the local decision making that shapes a Chapter 1 program. In her survey of local compensatory education directors, 29 percent reported receiving help from TACs, while 68 percent reported receiving technical assistance from their SEA. Turnbull reported a tendency for TAC services to be used more by larger districts, over 10,000 (35 percent) enrollment, than by smaller districts under 2,500 enrollment (23 percent). However the largest percentage of districts reporting as having used TAC services were medium sized with enrollment between 2,500 and 10,000 (37 percent). These figures may be deceiving, however. TAC experience with large districts has been in general very limited. For the most part the larger LEAs have their own evaluation staffs and have less need for assistance with data quality control, analysis and reporting than do smaller districts. When larger LEAs become involved with TAC it is usually on a project basis. That is the district identifies a need and seeks TAC assistance in developing a systematic solution to the need. One such example is the development of a longitudinal data base. Such a data base enables the LEA to track achievement patterns over time, following cohorts through the system. As program improvements are made their impact over a longer term can be observed. Differential growth patterns by school program component or subpopulation can also be observed, thus providing useful diagnostic information for targeting improvement efforts within the district.

The following examples illustrate some of the changes in Chapter 1 evaluation at the LEA level, and the role played by the TAC in the process. The first is from a very large district, in which the TAC has been working for several years on a number of special projects designed to strengthen the operation of the Chapter 1 program. The most notable of the projects are in the areas of data base development, quality assurance and testing.

"With TAC assistance, a longitudinal data base has been established for the Chapter 1 program. The data base is used for various purposes including the identification of effective teaching strategies and the monitoring of student progress.

As a followup to the longitudinal data base, a quality assurance process has been developed. In the process the LEA personnel work closely with school level staff to identify weaknesses in program operation and to institute changes for program improvement. The process generally begins with a review of program information and culminates in a program improvement plan through a series of onsite meetings and consultations. Preliminary results have been very encouraging.

TAC staff are presently working with the LEA staff to streamline the Chapter 1 testing program and coordinate it with other district testing activities. The district operates several compensatory education programs, funded by both federal and state monies. These programs often incur duplicated testing efforts which can be avoided if a common pool of test data is established."

Among medium to large size districts the development of data bases, using either a mainframe or microcomputer, has been a major priority area. Where these efforts have made use of microcomputers they have been based on commercially available data base programs such as DB Master, D Base II, Visifile, PFS and other similar packages. One of the important lessons learned by TAC in these efforts is that data base development is a long term activity, involving program staff, data processing staff and evaluation staff. Turnover among staff in the LEA or the TAC can deal a serious setback to such efforts. In the example cited above TAC involvement in the development and testing of the data base extended over nearly five years. In another district in Region 4 TAC has been working with the district on the development of a data base for nearly three years and the project is still not finished, due primarily to turnover among key personnel at the LEA.

District A provides an example of the progression that takes place when TAC works with an LEA to help computerize their Chapter 1 data analysis. This is a district of approximately 13,000 enrollment, with just over 800 Chapter 1 students. This district has progressed from using GE scores and less appropriate tests to now reporting in NCEs and using more appropriate and better administered tests. After experimenting with a variety of ways of computerizing their reporting they now have a TAC designed system using a microcomputer and a database management system. The evaluation is done by the Compensatory Education program unit. There

is a growing interest in developing a longitudinal data base which can be used for local program improvement purposes.

Smaller districts are also making good use of the new technology to develop student data bases and monitor quality control. District X, for example, has an enrollment of slightly under 3,000 and a Chapter 1 enrollment of less than 400. This LEA has progressed from a situation in which the evaluation responsibility was that of the building Title I teacher with no error screening, no consideration of norming dates, three different tests in use, no separation of pretest from posttest and all data were analyzed manually, to having a longitudinal database, uniformly appropriate testing procedures on a common achievement test selected by a district wide committee and a centralized evaluation. The availability of a microcomputer at the district offices has made the development and maintenance of the data base possible.

Summary and conclusions

Through the use of case studies this paper has illustrated the major types of change that have occurred in evaluating compensatory programs since the TIERS and TAC services were introduced in 1976. Major areas of change that have been discussed include testing practices, data quality, student selection, reporting and dissemination, applications of technology, and data use for program improvement.

Testing has improved in a variety of ways. Tests are being selected to better match the objectives and the needs of the Chapter 1 programs; knowledge of appropriate use of various metrics is greatly improved; tests are being administered with greater concern for norming dates and appropriate levels, all with the effect of improving the quality and usefulness of the data.

A logical progression from improved testing practices to better, more useful data has been illustrated. Improvements in testing however do not guarantee data of sufficient quality to be useful in program, district or state decision making. While issues of validity and reliability may undergird the testing process, the quality and ultimate usefulness of the data rest on the more mundane issue of human errors in the coding, transcribing, analyzing and reporting processes. For this reason systematic efforts to reduce error and assure data quality have been the focus of much of the TAC effort, and no doubt they will continue to be in the future. As we have seen, turnover in SEA or LEA staff during the course of the development of a quality control process can impede implementation. The approaches to quality control have been discussed in detail elsewhere. Examples discussed here have run the gamut from manually checking the data to developing sophisticated computer programs to perform edits. The method used in a particular SEA will depend on the availability of a computer at the SEA, the availability of microcomputers in the LEAs, the availability of technical expertise among staff, the number of districts reporting, and a number of other factors. This diversity among SEAs, already well documented, is a fact that requires a flexible and adaptive approach in technical assistance.

Improvements in reporting procedures follow from the improvements in testing and data quality. Reporting, however, is not the end of the evaluation - program improvement process. It is only a stage. In order for evaluation to have any real value as a tool for change the evaluation data must be disseminated to the right people in a usable format. This is an area of great improvement, as noted by Alkin in the Kraisner et al. study. It is an area in which improvements continue to be made and where much room for further improvement remains. Microcomputers have had an impact in this area by providing a means of creating a variety of data displays and by bringing data base management approaches within reach of nearly everyone. Unfortunately this exciting new technology is changing so rapidly that systems become outmoded almost before the bugs get worked out, leaving the user feeling that he or she will never quite catch up with the trend.

The application of high quality evaluation data to program improvement is an area in which changes have been slower to emerge and more difficult to accomplish, since as we have seen, they generally follow from changes in other areas. It is, however, the logical culmination of the evaluation process; the very purpose of evaluation. This year (1984-85) the TACs were specifically charged by ED, for the first time, with the responsibility of helping programs improve their effectiveness. A minimum of 25 percent of the TAC effort within each region was to be allocated to this goal. While many of us who have been with the TAC system since it began have always had this as a goal, putting it in concrete terms as a contractual requirement, has stimulated the TACs to re-examine our approach. We have begun to more seriously study the literature on effective schooling and change processes and to seek new models for bringing about change. This work has led us to seek a broadening of the concept of evaluation to that of a more continuous, open ended process intended to inform and focus improvement efforts.

Implications for future improvements

As we have seen many changes have occurred in the evaluation of compensatory education and the TACs have played a major role in effecting these changes. Many more improvements are needed however, and the TACs will be expected to play an even stronger role in bringing about change even though TAC funding has been reduced. Testing practices have improved greatly, but as test publishers bring new editions to market LEAs will continue to need guidance in selection and proper use of achievement tests. The burgeoning technology of the microcomputer has placed a powerful tool at the disposal of the educator from the classroom to the boardroom. Rapid advances in this technology make yesterday's solutions seem obsolete, therefore continuous support is needed for the Chapter 1 programs to make effective use of this technology. The TACs have played a key role in that process and must continue to give high priority to providing such support through training, materials development and consultation.

Program improvement is the area in which the greatest change is expected to occur. As the TAC's work in this area expands we are faced with an

emerging problem in terms of the focus of change. TACs are by virtue of their funding source, limited to dealing with Chapter 1 programs, their staff and their problems. We are prohibited from working with non-Chapter 1 staff such as the regular classroom teacher. The effective schooling literature focuses primarily on the building as the unit of change and identifies the building principal as the key instructional leader. Effectiveness factors such as leadership, clear objectives, school climate, instructional time, effective use of reinforcement, etc. are all seen as operating at school building or classroom level. Chapter 1 programs vary in structure but generally encompass all buildings in a district. Substantial variations in instructional approaches may occur from building to building within the same program, thus creating problems for TAC in effecting change at the building level. The TACs are working hard to develop strategies to overcome this problem. One strategy might be to link with other service providers using a similar model, but focus our efforts on the program across all schools, leaving the school level work to other service providers, who are not so restricted. Another strategy might be to develop a multi-level model with both a building and district approach. With this type of model TAC could provide training for district coordinators who could then replicate the process. TAC might need to identify a pilot school to work with in districts where this is attempted. Any model for school or program improvement must draw from the ever-growing body of literature on organizational change as applied to education as well as from the research on effective schooling.

Another key problem area in program improvement is the coordination of Chapter 1 with other programs and services. Since Chapter 1 is supplemental it would be expected that the most effective Chapter 1 programs would be those in which the instruction provided reinforces and supports the regular school curriculum. However one of the most common areas of concern for Chapter 1 program directors and building coordinators is the lack of coordination between the Chapter 1 and regular school programs. Once again this points to a problem, since TACs are restrained from working with the regular program.

Large districts pose a particular challenge to TACs. They tend not to use TAC services as much as the mid-size districts, yet they account for the largest proportion of Chapter 1 students. The most promising approach to helping large districts bring about improvements in the evaluation of Chapter 1 seems to be through assisting them with projects they see a need for. In this approach the TAC provides a specific expertise at no cost, that is essential to the completion of the project. It is intensive in nature and of longer duration. Both of these features also characterize program improvement assistance. To be more effective in assisting the large districts TACs must become more familiar with the specific needs and agendas of the large LEAs as they pertain to the evaluation and improvement of Chapter 1 programs. Recent discussions with some large district directors of research and evaluation indicate that they communicate with one another and draw upon pooled expertise to accomplish their goals. They have well defined priorities in the areas of classroom assessment, teacher evaluation, data base development and other problems that TACs are knowledgeable about. We

need to become part of their pool of expertise so that we can improve our understanding of their unique needs and they can benefit from our depth of experience in improving Chapter 1 evaluation.

Another major sub-area of Chapter 1 evaluation in which further improvements are needed is special programs such as migrant education and neglected or delinquent programs. During the past year ED has focused increased attention on improving the evaluation of these programs. The problems are many; the solutions few. A national work group comprised of TAC staff from each of the four regions and ED staff has been working in the problem of developing methods of evaluating migrant programs. This group has examined issues related to aggregation of evaluation data, use of the migrant student records transfer system (MSRTS), differences between "home base" states and states in which the migrants work when in transit. The work group is in the process of developing a set of recommendations for SEAs on evaluating migrant programs.

As this paper is being presented the TACs are preparing to compete under a new procurement from ED, to provide technical assistance services to Chapter 1 programs. Some changes are expected in the work scope of the RFP. One of the most important of the changes we anticipate is that the emphasis on program improvement will be increased, with as much as fifty percent of the TAC level of effort allocated to this task. A year from now at AERA, some TAC staff will be reporting at AERA on changes that have occurred in the area of program improvement as a result of this increased priority. Effective models will be developed for overcoming the problems identified here and improvements will be made, just as they have been made in the testing, evaluation and reporting areas.

References

- Alkin, M. The Utility of Title I Evaluation, in Reisner, et al. Assessment of the Title I Evaluation and Reporting System. Washington, D.C.: U.S. Department of Education, April, 1982.
- Boruch, R. Changes in Evaluation Practice at Local and State Levels and the Possible Influence of TIERS, in Reisner, et al. Assessment of the Title I Evaluation and Reporting System. Washington, D. C.: U.S. Department of Education, April, 1982.
- Reisner, E., Alkin, M., Boruch, R., Linn, R., and Millman, J. Assessment of the Title I Evaluation and Reporting System. Washington, D.C.: U.S. Department of Education, April, 1982.
- Stonehill, R. and Anderson, J. An Evaluation of ESEA Title I-Program Operations and Educational Effects: A Report to Congress. Washington, D.C.: U.S. Department of Education, Office of Planning, Budgeting and Evaluation, March, 1982.
- Turnbull, B., Technical Assistance and Local Program Implementation, in Title I, ESEA, A Special Report From The District Practices Study. McLean. Va.: Advanced Technology, September, 1982.